RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/510,363_
Source: _PCT_
Date Processed by STIC: _10-16-04_

ENTERED

PCT

<pre>
RAW SEQUENCE LISTING   DATE: 10/16/2004
PATENT APPLICATION: US/10/510,363 TIME: 09:23:07

Input Set : A:\SCHN0033.ST25.txt
Output Set: N:\CRF4\10162004\J510363.raw
</pre>

```
   3 <110> APPLICANT: Fishel, Richard A.
   4        Yoder, Kristine E.
   6 <120> TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT MODULATE A DNA REPAIR
   7        PATHWAY AND/OR RETRO-VIRAL INFECTIVITY, THE COMPOUNDS, AND USES
   8        THEREOF
  10 <130> FILE REFERENCE: SCHN-0033
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/510,363
C--> 12 <141> CURRENT FILING DATE: 2004-10-05
  12 <150> PRIOR APPLICATION NUMBER: PCT/US03/10302
  13 <151> PRIOR FILING DATE: 2003-04-04
  15 <150> PRIOR APPLICATION NUMBER: US 60/370,376
  16 <151> PRIOR FILING DATE: 2002-04-05
  18 <160> NUMBER OF SEQ ID NOS: 6
  20 <170> SOFTWARE: PatentIn version 3.3
  22 <210> SEQ ID NO: 1
  23 <211> LENGTH: 2751
  24 <212> TYPE: DNA
  25 <213> ORGANISM: Homo sapiens
  27 <400> SEQUENCE: 1
  28 gggagcttcc ggattgagcc ggaagtcccc ccagagcgga tgccgcggcg ggcctgtggg      60
  30 agcggggtca tcttctctct gctgctgtag ctgccatggg caaaagagac cgagcggacc     120
  32 gcgacaagaa gaaatccagg aagcggcact atgaggatga agaggatgat gaagaggacg     180
  34 ccccggggaa cgaccctcag gaagcggttc cctcggcggc ggggaagcag gtggatgagt     240
  36 caggcaccaa agtggatgaa tatggagcca aggactacag gctgcaaatg ccgctgaagg     300
  38 acgaccacac ctccaggccc ctctgggtgg ctcccgatgg ccatatcttc ttggaagcct     360
  40 tctctccagt ttacaaatat gcccaagact tcttggtggc tattgcagag ccagtgtgcc     420
  42 gaccaaccca tgtgcatgag tacaaactaa ctgcctactc cttgtatgca gctgtcagcg     480
  44 ttgggctgca aaccagtgac atcaccgagt acctcaggaa gctcagcaag actggagtcc     540
  46 ctgatggaat tatgcagttt attaagttgt gtactgtcag ctatggaaaa gtcaagctgg     600
  48 tcttgaagca aacagatac ttcgttgaaa gttgccaccc tgatgtaatc cagcatcttc     660
  50 tccaggaccc cgtgatccga gaatgccgct taagaaactc tgaaggggag gccactgagc     720
  52 tcatcacaga gactttcaca agcaaatctg ccatttctaa gactgctgaa agcagtggtg     780
  54 ggcccctccac ttcccgagtg acagatccac agggtaaatc tgacatcccc atggacctgt     840
  56 ttgacttcta tgagcaaatg gacaaggatg aagaagaaga agaagagaca cagacagtgt     900
  58 cttttgaagt caagcaggaa atgattgagg aactccagaa acgttgcatc cacctggagt     960
  60 accctctgtt ggcagaatat gacttccgga atgattctgt caaccctgat atcaacattg    1020
  62 acctaaagcc cacagctgtc ctcagaccct atcaggagaa gagcttgcga aagatgtttg    1080
  64 gaaacgggcg tgcacgttcg ggggtcattg ttcttccctg cggtgctgga aagtccctgg    1140
  66 ttggtgtgac tgctgcatgc actgtcagaa aacgctgtct ggtgctgggc aactcagctg    1200
  68 tttctgtgga gcagtggaaa gcccagttca agatgtggtc caccattgac gacagccaga    1260
  70 tctgccggtt cacctccgat gccaaggaca gcccatcgg ctgctccgtt gccattagca    1320
  72 cctactccat gctgggccac accaccaaaa ggtcctggga ggccgagcga gtcatggagt    1380
  74 ggctcaagac ccaggagtgg ggcctcatga tcctggatga agtgcacacc ataccagcca    1440
```

RAW SEQUENCE LISTING   DATE: 10/16/2004
PATENT APPLICATION: **US/10/510,363**   TIME: 09:23:07

Input Set : **A:\SCHN0033.ST25.txt**
Output Set: **N:\CRF4\10162004\J510363.raw**

```
 76 agatgttccg aagggtgctc accatcgtgc aggcccactg taagctgggt ttgactgcga  1500
 78 ccctcgtccg cgaagatgac aaaattgtgg atttaaattt tctgattggg cctaagctct  1560
 80 acgaagccaa ctggatggag ctgcagaata atggctacat cgccaaagtc cagtgtgctg  1620
 82 aggtctggtg ccctatgtct cctgaatttt accgggaata tgtggcaatc aaaaccaaga  1680
 84 aacgaatctt gctgtacacc atgaacccca acaaatttag agcttgccag tttctgatca  1740
 86 agtttcatga aaggaggaat gacaagatta ttgtctttgc tgacaatgtg tttgccctaa  1800
 88 aggaatatgc cattcgactg aacaaaccct atatctacgg acctacgtct caggggggaaa 1860
 90 ggatgcaaat tctccagaat ttcaagcaca accccaaaat taacaccatc ttcatatcca  1920
 92 aggtaggtga cacttcgttt gatctgccgg aagcaaatgt cctcattcag atctcatccc  1980
 94 atggtggctc caggcgtcag gaagcccaaa ggctagggcg ggtgcttcga gctaaaaaag  2040
 96 ggatggttgc agaagagtac aatgcctttt tctactcact ggtatcccag gacacacagg  2100
 98 aaatggctta ctcaaccaag cggcagagat tcttggtaga tcaaggttat agcttcaagg  2160
100 tgatcacgaa actcgctggc atggaggagg aagacttggc gttttcgaca aaagaagagc  2220
102 aacagcagct cttacagaaa gtcctggcag ccactgacct ggatgccgag gaggaggtgg  2280
104 tggctgggga atttggctcc agatccagcc aggcatctcg gcgctttggc accatgagtt  2340
106 ctatgtctgg ggccgacgac actgtgtaca tggagtacca ctcatcgcgg agcaaggcgc  2400
108 ccagcaaaca tgtacacccg ctcttcaagc gctttaggaa atgatgctta ggcagggtac  2460
110 ttcgttcaag accggcgctt ggcacccttg ttggaaaggg attttcagca taacattttc  2520
112 cttccacctc tttgaccttc cctccagcgt tggccaaatt gtgctgagga agatgcatca  2580
114 agggcttggc tgtgccttca taggtcatct agggtttttat aaaggaggag gagacaatat  2640
116 tttttcaaac ttttttgggga gtggggtcat ttctgtatat aaaaaatgtt aatatttaag  2700
118 gtgtatttat gttaccgttc tgaataaaca gaatggacca ttgaaccagt a            2751
121 <210> SEQ ID NO: 2
122 <211> LENGTH: 782
123 <212> TYPE: PRT
124 <213> ORGANISM: Homo sapiens
126 <400> SEQUENCE: 2
128 Met Gly Lys Arg Asp Arg Ala Asp Arg Asp Lys Lys Lys Ser Arg Lys
129 1               5               10              15
132 Arg His Tyr Glu Asp Glu Glu Asp Asp Glu Glu Asp Ala Pro Gly Asn
133           20              25              30
136 Asp Pro Gln Glu Ala Val Pro Ser Ala Ala Gly Lys Gln Val Asp Glu
137           35              40              45
140 Ser Gly Thr Lys Val Asp Glu Tyr Gly Ala Lys Asp Tyr Arg Leu Gln
141     50              55              60
144 Met Pro Leu Lys Asp Asp His Thr Ser Arg Pro Leu Trp Val Ala Pro
145 65              70              75              80
148 Asp Gly His Ile Phe Leu Glu Ala Phe Ser Pro Val Tyr Lys Tyr Ala
149           85              90              95
152 Gln Asp Phe Leu Val Ala Ile Ala Glu Pro Val Cys Arg Pro Thr His
153           100             105             110
156 Val His Glu Tyr Lys Leu Thr Ala Tyr Ser Leu Tyr Ala Ala Val Ser
157           115             120             125
160 Val Gly Leu Gln Thr Ser Asp Ile Thr Glu Tyr Leu Arg Lys Leu Ser
161     130             135             140
164 Lys Thr Gly Val Pro Asp Gly Ile Met Gln Phe Ile Lys Leu Cys Thr
165 145             150             155             160
168 Val Ser Tyr Gly Lys Val Lys Leu Val Leu Lys His Asn Arg Tyr Phe
169           165             170             175
```

```
172 Val Glu Ser Cys His Pro Asp Val Ile Gln His Leu Leu Gln Asp Pro
173             180              185              190
176 Val Ile Arg Glu Cys Arg Leu Arg Asn Ser Glu Gly Glu Ala Thr Glu
177         195              200              205
180 Leu Ile Thr Glu Thr Phe Thr Ser Lys Ser Ala Ile Ser Lys Thr Ala
181         210              215              220
184 Glu Ser Ser Gly Gly Pro Ser Thr Ser Arg Val Thr Asp Pro Gln Gly
185 225             230              235                  240
188 Lys Ser Asp Ile Pro Met Asp Leu Phe Asp Phe Tyr Glu Gln Met Asp
189             245              250                  255
192 Lys Asp Glu Glu Glu Glu Glu Glu Thr Gln Thr Val Ser Phe Glu Val
193         260              265              270
196 Lys Gln Glu Met Ile Glu Glu Leu Gln Lys Arg Cys Ile His Leu Glu
197         275              280              285
200 Tyr Pro Leu Leu Ala Glu Tyr Asp Phe Arg Asn Asp Ser Val Asn Pro
201     290              295              300
204 Asp Ile Asn Ile Asp Leu Lys Pro Thr Ala Val Leu Arg Pro Tyr Gln
205 305             310              315                  320
208 Glu Lys Ser Leu Arg Lys Met Phe Gly Asn Gly Arg Ala Arg Ser Gly
209             325              330              335
212 Val Ile Val Leu Pro Cys Gly Ala Gly Lys Ser Leu Val Gly Val Thr
213             340              345              350
216 Ala Ala Cys Thr Val Arg Lys Arg Cys Leu Val Leu Gly Asn Ser Ala
217         355              360              365
220 Val Ser Val Glu Gln Trp Lys Ala Gln Phe Lys Met Trp Ser Thr Ile
221     370          .      375              380
224 Asp Asp Ser Gln Ile Cys Arg Phe Thr Ser Asp Ala Lys Asp Lys Pro
225 385             390              395                  400
228 Ile Gly Cys Ser Val Ala Ile Ser Thr Tyr Ser Met Leu Gly His Thr
229             405              410              415
232 Thr Lys Arg Ser Trp Glu Ala Glu Arg Val Met Glu Trp Leu Lys Thr
233             420              425              430
236 Gln Glu Trp Gly Leu Met Ile Leu Asp Glu Val His Thr Ile Pro Ala
237         435              440              445
240 Lys Met Phe Arg Arg Val Leu Thr Ile Val Gln Ala His Cys Lys Leu
241     450              455              460
244 Gly Leu Thr Ala Thr Leu Val Arg Glu Asp Asp Lys Ile Val Asp Leu
245 465             470              475                  480
248 Asn Phe Leu Ile Gly Pro Lys Leu Tyr Glu Ala Asn Trp Met Glu Leu
249             485              490              495
252 Gln Asn Asn Gly Tyr Ile Ala Lys Val Gln Cys Ala Glu Val Trp Cys
253             500              505              510
256 Pro Met Ser Pro Glu Phe Tyr Arg Glu Tyr Val Ala Ile Lys Thr Lys
257         515              520              525
260 Lys Arg Ile Leu Leu Tyr Thr Met Asn Pro Asn Lys Phe Arg Ala Cys
261     530              535              540
264 Gln Phe Leu Ile Lys Phe His Glu Arg Arg Asn Asp Lys Ile Ile Val
265 545             550              555                  560
268 Phe Ala Asp Asn Val Phe Ala Leu Lys Glu Tyr Ala Ile Arg Leu Asn
```

```
269                565                570                575
272 Lys Pro Tyr Ile Tyr Gly Pro Thr Ser Gln Gly Glu Arg Met Gln Ile
273                580                585                590
276 Leu Gln Asn Phe Lys His Asn Pro Lys Ile Asn Thr Ile Phe Ile Ser
277            595                600                605
280 Lys Val Gly Asp Thr Ser Phe Asp Leu Pro Glu Ala Asn Val Leu Ile
281        610                615                620
284 Gln Ile Ser Ser His Gly Gly Ser Arg Arg Gln Glu Ala Gln Arg Leu
285 625                630                635                640
288 Gly Arg Val Leu Arg Ala Lys Lys Gly Met Val Ala Glu Glu Tyr Asn
289                645                650                655
292 Ala Phe Phe Tyr Ser Leu Val Ser Gln Asp Thr Gln Glu Met Ala Tyr
293            660                665                670
296 Ser Thr Lys Arg Gln Arg Phe Leu Val Asp Gln Gly Tyr Ser Phe Lys
297        675                680                685
300 Val Ile Thr Lys Leu Ala Gly Met Glu Glu Glu Asp Leu Ala Phe Ser
301        690                695                700
304 Thr Lys Glu Glu Gln Gln Gln Leu Leu Gln Lys Val Leu Ala Ala Thr
305 705                710                715                720
308 Asp Leu Asp Ala Glu Glu Glu Val Val Ala Gly Glu Phe Gly Ser Arg
309                725                730                735
312 Ser Ser Gln Ala Ser Arg Arg Phe Gly Thr Met Ser Ser Met Ser Gly
313            740                745                750
316 Ala Asp Asp Thr Val Tyr Met Glu Tyr His Ser Ser Arg Ser Lys Ala
317        755                760                765
320 Pro Ser Lys His Val His Pro Leu Phe Lys Arg Phe Arg Lys
321        770                775                780
324 <210> SEQ ID NO: 3
325 <211> LENGTH: 2318
326 <212> TYPE: DNA
327 <213> ORGANISM: Homo sapiens
329 <400> SEQUENCE: 3
330 atgaagctca acgtggacgg gctcctggtc tacttcccgt acgactacat ctaccccgag      60
332 cagttctcct acatgcggga gctcaaacgc acgctggacg ccaagggtca tggagtcctg     120
334 gagatgccct caggcaccgg gaagacagta tccctgttgg ccctgatcat ggcataccag     180
336 agagcatatc cgctggaggt gaccaaactc atctactgct caagaactgt gccagagatt     240
338 gagaaggtga ttgaagagct tcgaaagttg ctcaacttct atgagaagca ggagggcgag     300
340 aagctgccgt ttctgggact ggctctgagc tcccgcaaaa acttgtgtat tcaccctgag     360
342 gtgacacccc tgcgctttgg gaaggacgtc gatgggaaat gccacagcct cacagcctcc     420
344 tatgtgcggg cgcagtacca gcatgacacc agcctgcccc actgccgatt ctatgaggaa     480
346 tttgatgccc atgggcgtga ggtgcccctc cccgctggca tctacaacct ggatgacctg     540
348 aaggccctgg ggcggcgcca gggctggtgc ccatacttcc ttgctcgata ctcaatcctg     600
350 catgccaatg tggtggttta tagctaccac tacctcctgg accccaagat tgcagacctg     660
352 gtgtccaagg aactggcccg caaggccgtc gtggtcttcg acgaggccca caacattgac     720
354 aacgtctgca tcgactccat gagcgtcaac ctcacccgcc ggacccttga ccggtgccag     780
356 ggcaacctgg agaccctgca agacggtg ctcaggatca aagagacaga cgagcagcgc     840
358 ctgcgggacg agtaccggcg tctggtggag gggctgcggg aggccagcgc cgcccgggag     900
360 acggacgccc acctggccaa ccccgtgctg cccgacgaag tgctgcagga ggcagtgcct     960
362 ggctccatcc gcacggccga gcatttcctg ggcttcctga ggcggctgct ggagtacgtg    1020
```

RAW SEQUENCE LISTING   DATE: 10/16/2004
PATENT APPLICATION: US/10/510,363   TIME: 09:23:07

Input Set : A:\SCHN0033.ST25.txt
Output Set: N:\CRF4\10162004\J510363.raw

```
364 aagtggcggc tgcgtgtgca gcatgtggtg caggagagcc cgcccgcctt cctgagcggc  1080
366 ctggcccagc gcgtgtgcat ccagcgcaag cccctcagat tctgtgctga acgcctccgg  1140
368 tccctgctgc atactctgga gatcaccgac cttgctgact tctccccgct cacccctcctt  1200
370 gctaactttg ccaccccttgt cagcacctac gccaaaggct tcaccatcat catcgagccc  1260
372 tttgacgaca gaaccccgac cattgccaac cccatcctgc acttcagctg catggacgcc  1320
374 tcgctggcca tcaaacccgt atttgagcgt ttccagtctg tcatcatcac atctgggaca  1380
376 ctgtccccgc tggacatcta ccccaagatc ctggacttcc accccgtcac catggcaacc  1440
378 ttcaccatga cgctggcacg ggtctgcctc tgccctatga tcatcggccg tggcaatgac  1500
380 caggtggcca tcagctccaa atttgagacc cgggaggata ttgctgtgat ccggaactat  1560
382 gggaacctcc tgctggagat gtccgctgtg gtccctgatg gcatcgtggc cttcttcacc  1620
384 agctaccagt acatggagag caccgtggcc tcctggtatg agcaggggat ccttgagaac  1680
386 atccagagga caaagctgct ctttattgag acccaggatg gtgccgaaac cagtgtcgcc  1740
388 ctggagaagt accaggaggc ctgcgagaat ggccgcgggg ccatcctgct gtcagtggcc  1800
390 cggggcaaag tgtccgaggg aatcgacttt gtgcaccact acgggcgggc cgtcatcatg  1860
392 tttggcgtcc cctacgtcta cacacagagc cgcattctca aggcgcggct ggaatacctg  1920
394 cgggaccagt tccagattcg tgagaatgac tttcttacct tcgatgccat gcgccacgcg  1980
396 gcccagtgtg tgggtcgggc catcaggggc aagacggact acggcctcat ggtctttgcc  2040
398 gacaagcggt ttgcccgtgg ggacaagcgg gggaagctgc cccgctggat ccaggagcac  2100
400 ctcacagatg ccaacctcaa cctgaccgtg gacgagggtg tccaggtggc caagtacttc  2160
402 ctgcggcaga tggcacagcc cttccaccgg gaggatcagc tgggcctgtc cctgctcagc  2220
404 ctggagcagc tagaatcaga ggagacgctg aagaggatag agcagattgc tcagcagctc  2280
406 tgagtggggc gggtgggggcc ataaacggtt cctggtga                           2318
409 <210> SEQ ID NO: 4
410 <211> LENGTH: 760
411 <212> TYPE: PRT
412 <213> ORGANISM: Homo sapiens
414 <400> SEQUENCE: 4
416 Met Lys Leu Asn Val Asp Gly Leu Leu Val Tyr Phe Pro Tyr Asp Tyr
417 1               5                   10                  15
420 Ile Tyr Pro Glu Gln Phe Ser Tyr Met Arg Glu Leu Lys Arg Thr Leu
421             20                  25                  30
424 Asp Ala Lys Gly His Gly Val Leu Glu Met Pro Ser Gly Thr Gly Lys
425         35                  40                  45
428 Thr Val Ser Leu Leu Ala Leu Ile Met Ala Tyr Gln Arg Ala Tyr Pro
429     50                  55                  60
432 Leu Glu Val Thr Lys Leu Ile Tyr Cys Ser Arg Thr Val Pro Glu Ile
433 65                  70                  75                  80
436 Glu Lys Val Ile Glu Glu Leu Arg Lys Leu Leu Asn Phe Tyr Glu Lys
437             85                  90                  95
440 Gln Glu Gly Glu Lys Leu Pro Phe Leu Gly Leu Ala Leu Ser Ser Arg
441             100                 105                 110
444 Lys Asn Leu Cys Ile His Pro Glu Val Thr Pro Leu Arg Phe Gly Lys
445         115                 120                 125
448 Asp Val Asp Gly Lys Cys His Ser Leu Thr Ala Ser Tyr Val Arg Ala
449     130                 135                 140
452 Gln Tyr Gln His Asp Thr Ser Leu Pro His Cys Arg Phe Tyr Glu Glu
453 145                 150                 155                 160
456 Phe Asp Ala His Gly Arg Glu Val Pro Leu Pro Ala Gly Ile Tyr Asn
457             165                 170                 175
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 10/16/2004
PATENT APPLICATION: **US/10/510,363**   TIME: 09:23:08

Input Set : **A:\SCHN0033.ST25.txt**
Output Set: **N:\CRF4\10162004\J510363.raw**

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:5,6

VERIFICATION SUMMARY   DATE: 10/16/2004
PATENT APPLICATION: **US/10/510,363**   TIME: 09:23:08

Input Set : **A:\SCHN0033.ST25.txt**
Output Set: **N:\CRF4\10162004\J510363.raw**

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date